

REPORT CARDS & ACCOUNTABILITY SYSTEMS

BRADLEY CARL
VALUE-ADDED RESEARCH CENTER
UW-MADISON

April 10, 2013

Potential Purposes/Goals of Report Cards & School Accountability Systems

- Provide information to the public about school performance and/or quality
- Reinforce the need for high expectations (and the pursuit of effective strategies for reaching them)
- Compare performance across schools that may serve very different student populations
- Make higher-stakes decisions (additional oversight, closures, funding, etc.)
 - ▣ Choices about which metrics to include in report cards, and how to weight them relative to one another, should be driven by the intended uses of the data

Key Policy Considerations

- Which measures are optimal for reinforcing which goals?
 - ▣ Attainment (point-in-time) metrics may be appropriate for reinforcing high expectations, but less useful for making fair comparisons across schools that serve different students
- What constraints (if any) are placed on growth models by policymakers (e.g., prohibiting the use of certain student demographic factors), and how can these constraints best be addressed?
- To what extent can/should the growth measures used in different accountability systems be aligned?
 - ▣ Implications of using one measure for school-level report cards and another for educator effectiveness decisions

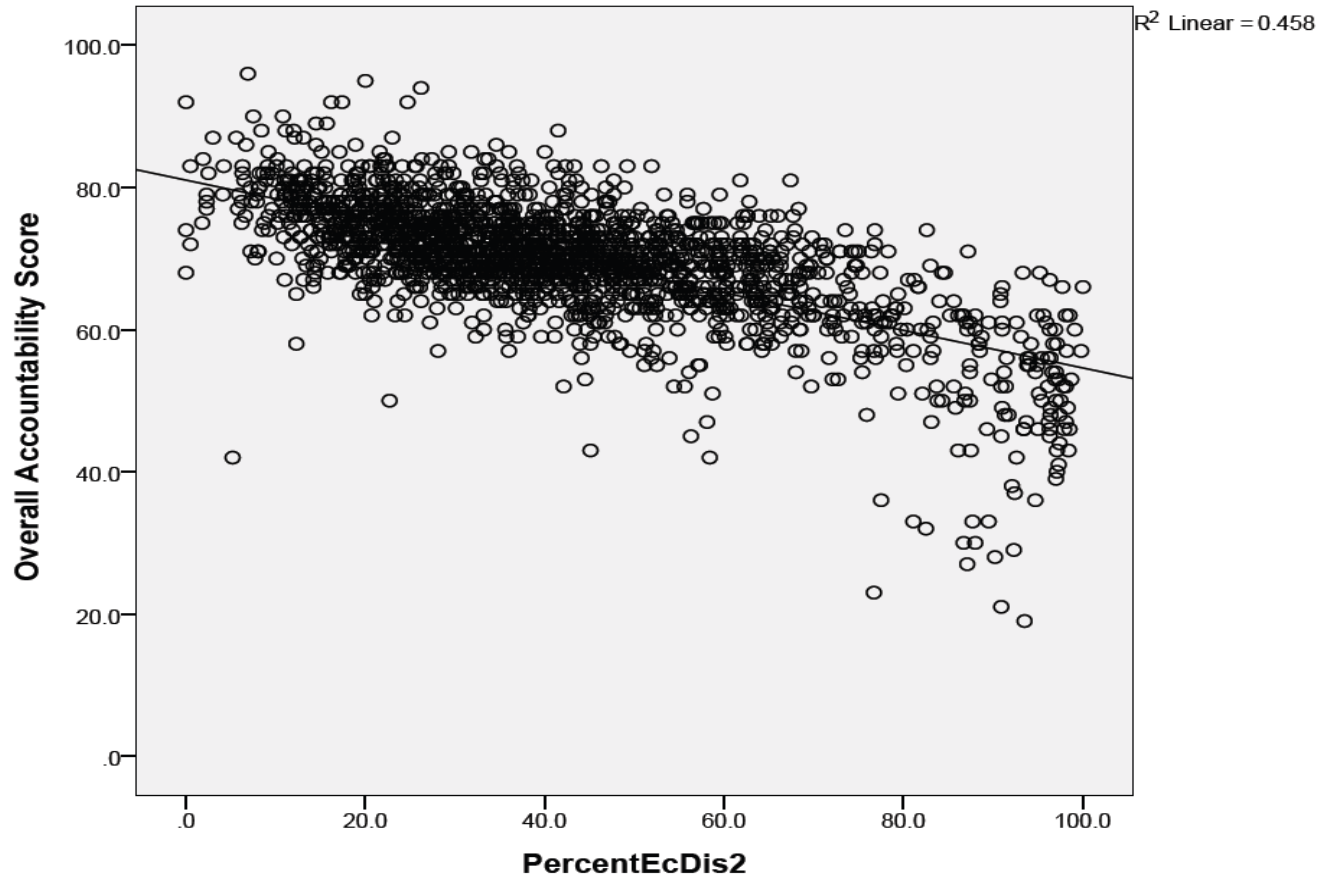
Key Policy Considerations (cont'd)

- Is the growth of all students treated equally (e.g., growth within proficiency categories as well as across categories), and are equivalent amounts of growth and decline treated equally?
- Does model address differences in growth at different achievement levels (including floor and ceiling) and due to test measurement error?
- How to strike the optimal balance between accuracy (technical validity), fairness, and complexity/transparency?

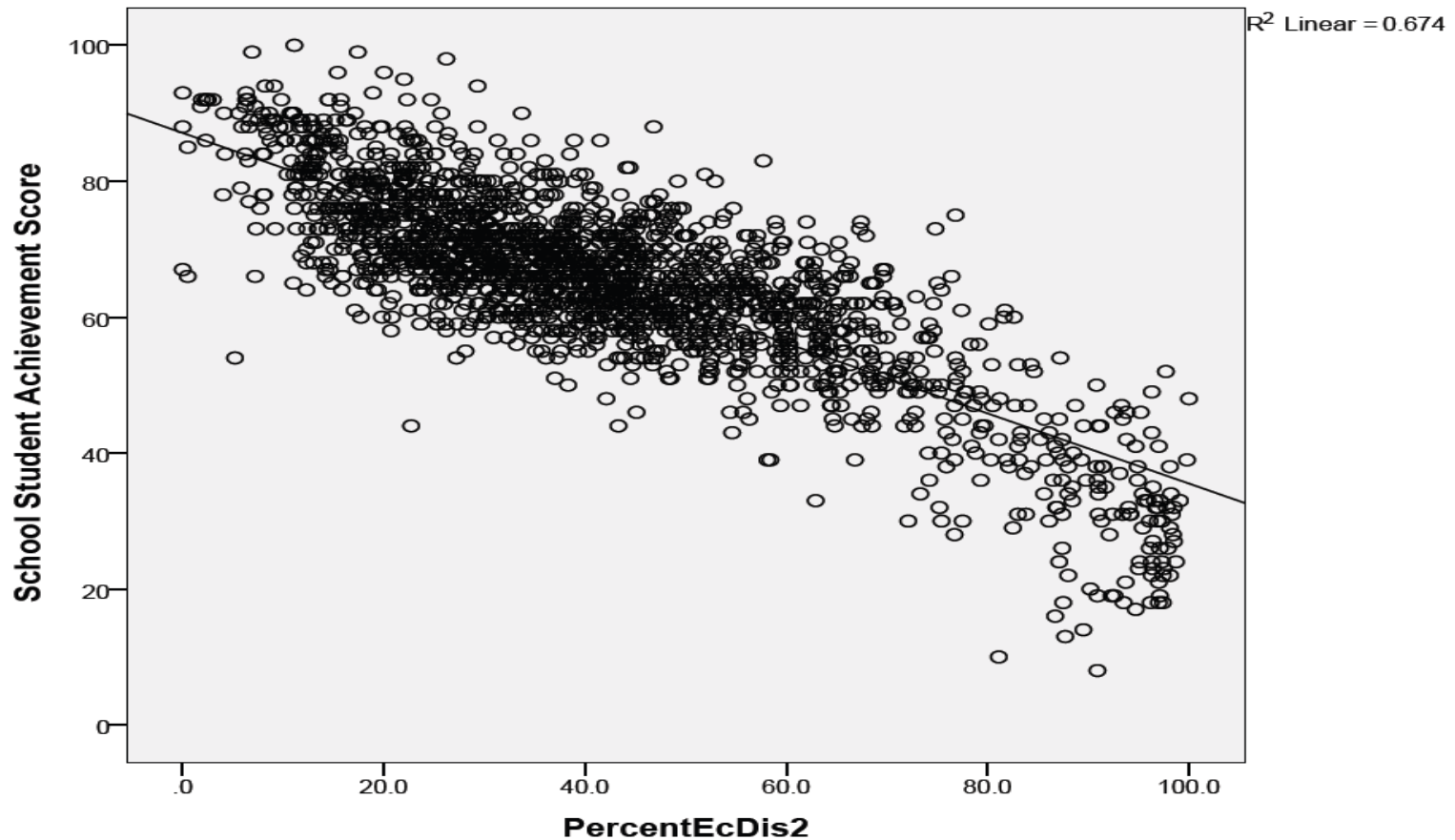
Key Policy Considerations (cont'd)

- How sensitive is/are the measure(s) to differences across schools around factors beyond their control (e.g., types of students served)?
- Policymakers should carefully consider the consequential validity (intended and unintended consequences) and incentives that accountability systems and growth models convey

Overall Index Score by Poverty Rate



Achievement Score by Poverty Rate



Different Ways to Measure Student Performance

- Many different ways exist to measure student performance (for Report Cards and in general):
 - ▣ Attainment: point in time/relative to standard (% proficient)
 - ▣ Simple gain: change in scale score or % proficient
 - ▣ Advanced growth models (student growth percentiles, value-added, etc.)

Comparing Growth Models

- Important to match choice of growth model with intended uses and level of stakes: public information, school-level or educator-level accountability, etc.
- Correlation between most advanced growth models is approximately 80-90%, which indicates that there are large differences for some schools and educators
- Differences across models are not neutral with respect to key student attributes
 - ▣ Districts/schools/teachers that serve larger populations of lower-income and non-white students will generally look worse when value-added is not used

Other Threats to Validity

- ❑ Small sample size (a big issue in Wisconsin): failure to account for this may result in small schools appearing to be the lowest-performing and highest-performing in the state
- ❑ Precision: avoid over-interpretation of data via “full intellectual honesty” (confidence intervals)

Growth Models in Report Cards

- Results from growth models typically represent only one component of Report Cards
- Not currently possible to measure growth for all grades and schools in WI (high schools, early elementary, non-reading/math), although upcoming changes to state testing system (starting 2014-15) may provide new opportunities
- Depending on purposes for which it is/will be used, a review of either the growth model component or the broader set of Report Card metrics may be appropriate